Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Network Theory
4.02

EEU45C09 / EEP55C09
Self Organising Technological Networks

CONNECT
Networks of the Future

Nicola Marchetti
nicola.marchetti@tcd.ie

Random Graphs: Edge probability p

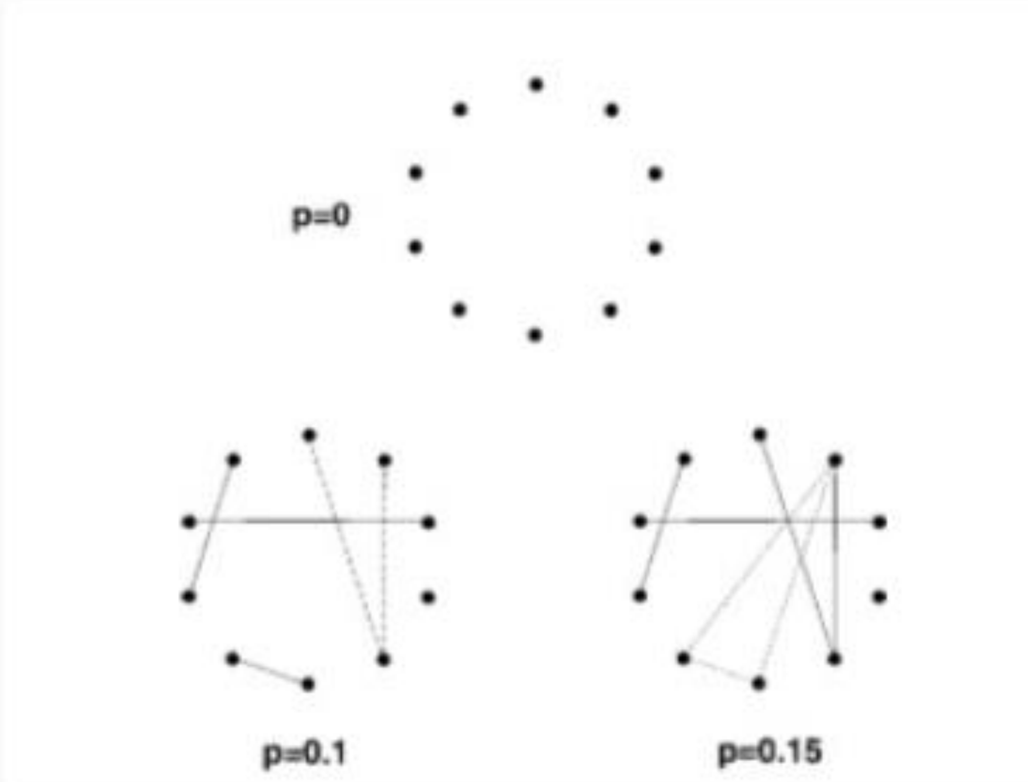

Figure 2: A picture of the resulting random graphs from evolutions with different p values. We note the emergence of trees and cycles in the $p = 0.1$ and $p = 0.15$ graphs respectively.

Random Graphs: Threshold probabilities

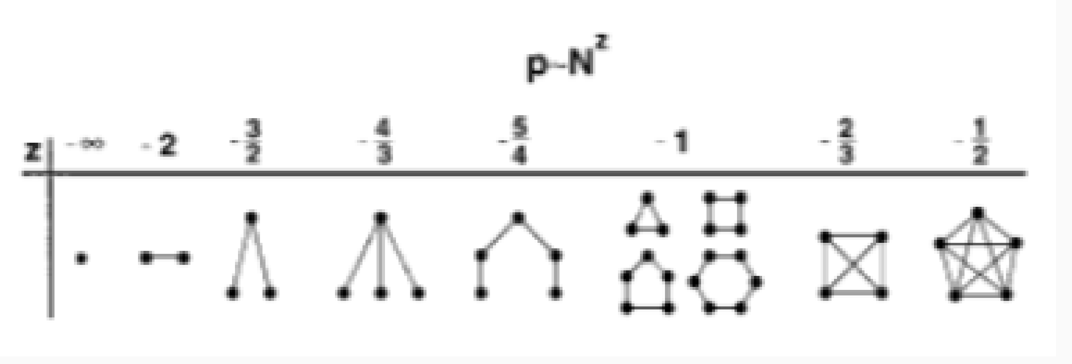

Figure 3: The threshold probabilities at which different subgraphs appear in a random graph. For $p \sim N^{-3/2}$, trees of order 3 appear, while for $p \sim N^{-4/3}$ trees of order 4 appear. At $p \sim N^{-1}$, trees of all orders and cycles appear.

Random Graphs: A comparison to real-world networks

TABLE I. The general characteristics of several real networks. For each network we have indicated the number of nodes, the average degree $\langle k \rangle$, the average path length ℓ, and the clustering coefficient C. For a comparison we have included the average path length ℓ_{rand} and clustering coefficient C_{rand} of a random graph of the same size and average degree. The numbers in the last column are keyed to the symbols in Figs. 8 and 9.

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference	Nr.
WWW, site level, undir.	153 127	35.21	3.1	3.35	0.1078	0.00023	Adamic, 1999	1
Internet, domain level	3015–6209	3.52–4.11	3.7–3.76	6.36–6.18	0.18–0.3	0.001	Yook *et al.*, 2001a, Pastor-Satorras *et al.*, 2001	2
Movie actors	225 226	61	3.65	2.99	0.79	0.00027	Watts and Strogatz, 1998	3
LANL co-authorship	52 909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman, 2001a, 2001b, 2001c	4
MEDLINE co-authorship	1 520 251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman, 2001a, 2001b, 2001c	5
SPIRES co-authorship	56 627	173	4.0	2.12	0.726	0.003	Newman, 2001a, 2001b, 2001c	6
NCSTRL co-authorship	11 994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman, 2001a, 2001b, 2001c	7
Math. co-authorship	70 975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.*, 2001	8
Neurosci. co-authorship	209 293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási *et al.*, 2001	9
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner and Fell, 2000	10
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner and Fell, 2000	11
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya and Solé, 2000	12
Silwood Park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya and Solé, 2000	13
Words, co-occurrence	460.902	70.13	2.67	3.03	0.437	0.0001	Ferrer i Cancho and Solé, 2001	14
Words, synonyms	22 311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.*, 2001b	15
Power grid	4941	2.67	18.7	12.4	0.08	0.005	Watts and Strogatz, 1998	16
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts and Strogatz, 1998	17

Figure 4: Many real-world networks have a similar average path length l, but larger clustering coefficient C to random graphs.

Small World Experiment

Stanley Milgram: Six degrees of separation [1]

$N = 296$ starting subjects asked to send letters from Boston and Nebraska to Massachusetts

64 Chains reach the target person

The mean number of intermediaries between source and target is 5.2

48% of letter pass through three people (so-called “stars”)

Small-World Experiment: Procedure

An arbitrary "target-person" and a group of "starting-persons" were selected.

An attempt was made to generate an acquaintance chain from each "starter" to each "target". A document was given to each starter with the name of the target person. The document stipulated that the starter may only send the document to a "first-name" basis acquaintance. The document contained some information about the "target" which was used to advance the chain (eg. professional information). Data about each participant was gathered eg. age, sex, profession.

Small-World Experiment: Results

217 of 296 starting persons actually sent the document on to friends.

29% completion rate. 64 documents reached the target

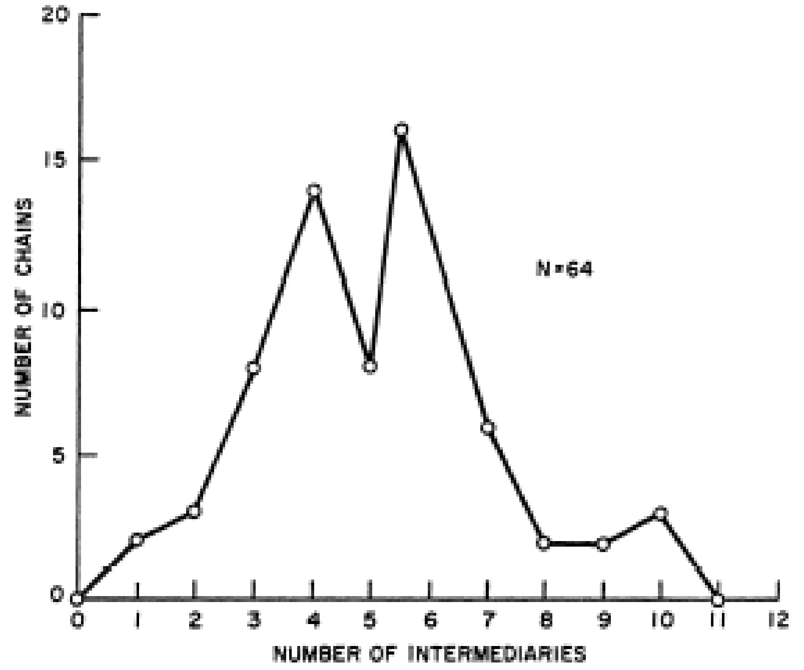

FIGURE 1

Lengths of Completed Chains

Small-World Networks

The first successful attempt to generate graphs with high clustering coefficients and small average path length is that of Watts and Strogatz [2].

Watts-Strogatz Model

Generate random graphs with a short average path length and a high clustering coefficient! To construct one of these graphs:

1. Begin with N nodes attached in a ring-lattice. Every node is connected to its K neighbours $(K/2)$ on each side.
2. Randomly rewire each edge of the ring with probability p such that self-connections and duplicate edges are avoided.

This process introduces $pNK/2$ long-range edges connecting nodes from different neighbourhoods in the graph.

Watts-Strogatz Model: Example

By varying p, we transition from an ordered ring (at $p = 0$) to a random graph (at $p = 1$).

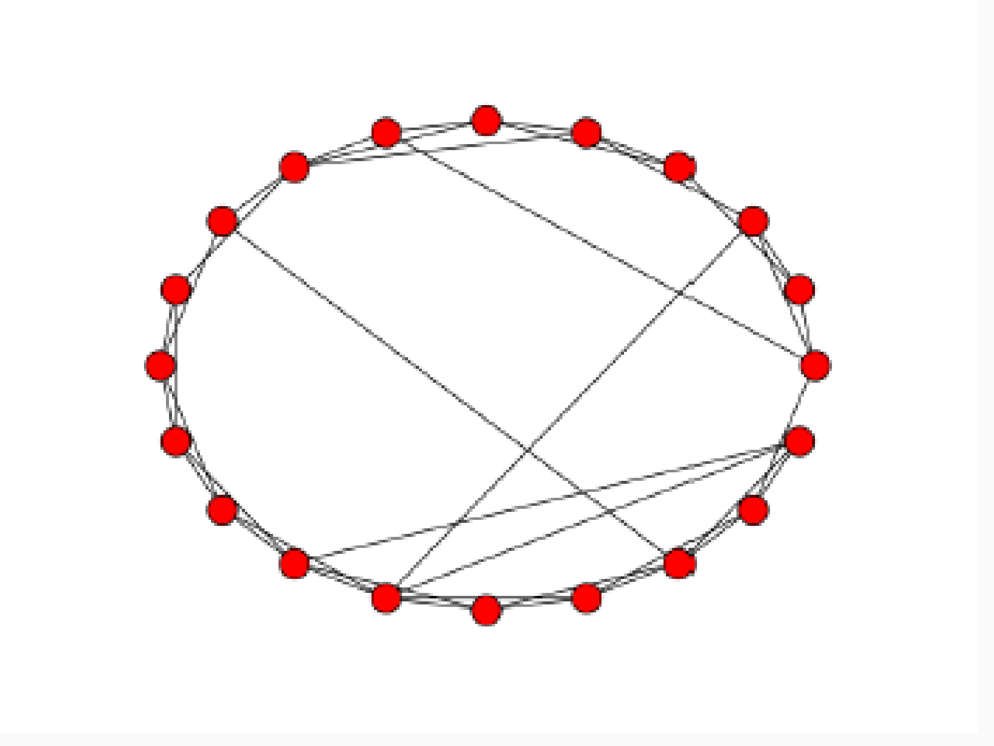

Figure 5: Watts-Strogatz model [2]: $|V| = 20$, $\langle k \rangle = 4$, $p = 0.2$

Watts-Strogatz Model: The effect of p

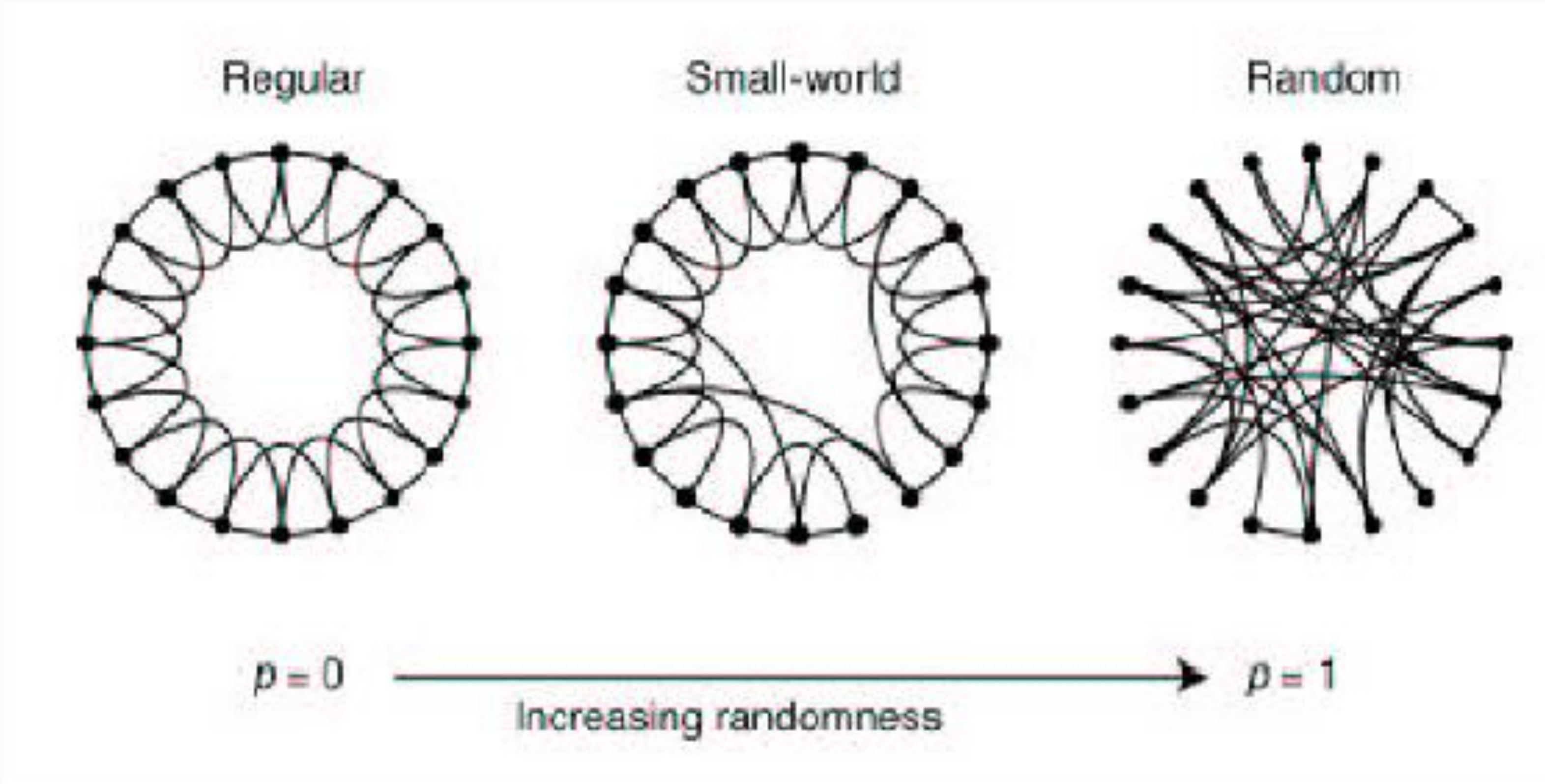

Figure 6: At very large and small values of p the Watts-Strogatz model resembles the random and regular graphs.

Watts-Strogatz Model: Small-World

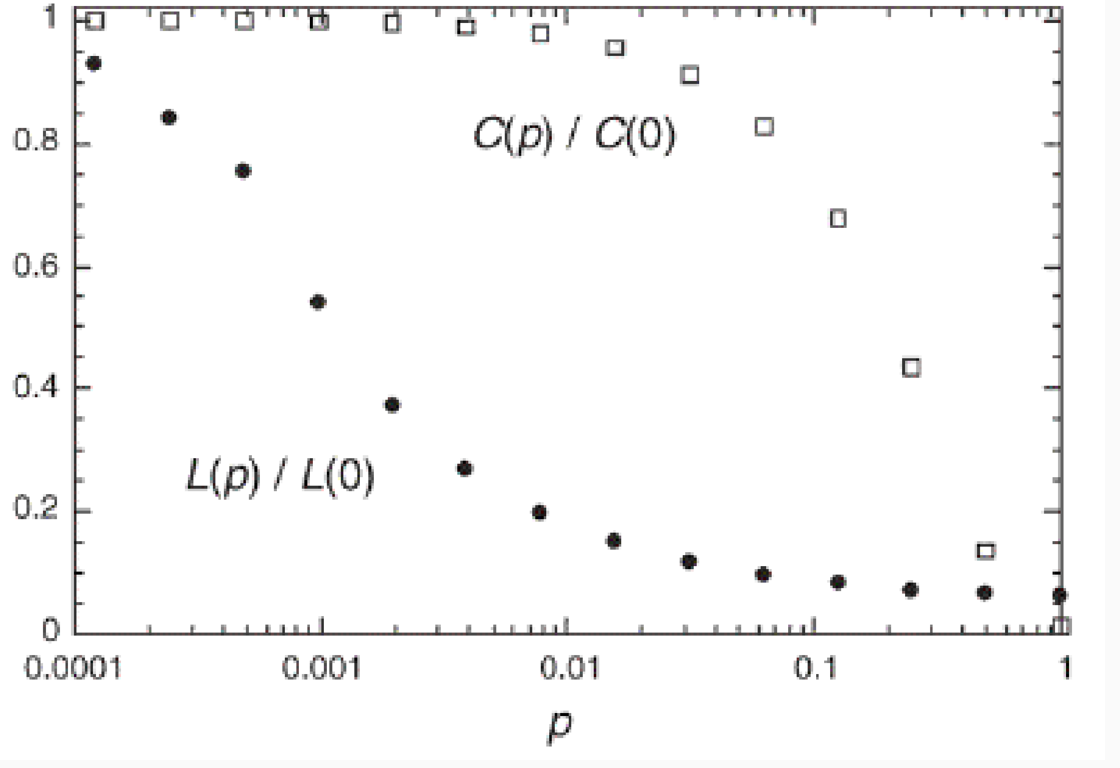

Figure 7: For intermediate values of p, the average path length is small and the clustering coefficient is large.

Watts-Strogatz model: Average Path Length

The Watts-Strogatz model has its roots in social systems.

It copies the behaviour of social networks:

- You are very likely to know your neighbours, colleagues and people your friends introduce you to (local friends).
- Everyone has some friends which are far-away (people in other countries, old acquaintances).
- These distant friends are represented by the long-range connections in the Watts-Strogatz Network.

Watts-Strogatz model

Let us consider the average path length $\langle l \rangle(p)$ and clustering coefficient $C(p)$ as a function of p.

For a ring-lattice, $\langle l \rangle(0) \approx N/2K \gg 1$ and $C(0) \approx 3/4$. Therefore $\langle l \rangle(0)$ scales linearly and $C(0)$ is large.

Watts-Strogatz model

On the other hand, as $p \to 1$, the model converges to a random graph.

From lecture 1, $\langle l \rangle(1) \sim \log N / \log K$ and $C(1) \sim K/N$.

These limiting values suggest that large C is always associated with large $\langle l \rangle$ and vice versa, however as we have seen in Fig. 7 there is a broad interval of p over which $\langle l \rangle(p)$ is close to $\langle l \rangle(1)$ yet $C(p) \gg C(1)$.

Watts-Strogatz model: Average Path Length

As p increases, there is a change in the scaling of the average path length $\langle l \rangle$.

At low p, $\langle l \rangle$ scales linearly with the graph size, while for large p the scaling is logarithmic.

The initial rapid drop in $\langle l \rangle$ stems from the appearance of shortcuts through the graph. Every shortcut is likely to connect well separated areas of the graph.

Watts-Strogatz model: Average Path Length

Does the onset of small-world behaviour (small $\langle l \rangle$ and large C) depend on the graph size?

$\langle l \rangle$ does not begin to decrease until $p \geq 2/NK$, guaranteeing the existence of at least one shortcut.

This implies there is a p-dependent crossover-size of graphs N^* such that if $N < N^*, \langle l \rangle \sim N$ but if $N > N^*, \langle l \rangle \sim \log N$.

Watts-Strogatz model: Average Path Length

The characteristic path length scales like,

$$\langle l \rangle (N, p) \sim N^* F\left(\frac{N}{N^*}\right),$$

where,

$$F(u) = \begin{cases} u, & \text{if } u \ll 1 \\ \log u, & \text{if} u \gg 1. \end{cases}$$

Watts-Strogatz model: Average Path Length

Using numerical simulations and analytic arguments, it has been shown that the crossover size N^* scales with p as $N^* \sim p^{-1/d}$, where d is the dimension of the original lattice.

In our case, the one dimensional ring leads to the crossover $N^* \sim p^{-1}$.

This implies that the small world behaviour occurs at a critical rewiring probability,

$$p^* \sim 1/N.$$

Watts-Strogatz model: Clustering coefficient

We now turn to the clustering coefficient of the WS model. What is the dependence of $C(p)$ on p?

The clustering coefficient is the fraction of the mean number of edges between the neighbours of a node and the mean number of possible edges between those neighbours.

Watts-Strogatz model: Clustering coefficient

To calculate $C(p)$ for the WS model, start with a regular lattice with a clustering coefficient of $C(0)$.

For $p > 0$, two neighbours of a node i that were connected in the original lattice are still neighbours of i and connected by an edge with probability $(1-p)^3$, since there are 3 edges which must remain in place. Therefore,

$$C(p) \approx C(0)(1-p)^3.$$

(There is a contribution to the clustering coefficient from rewired edges but this is vanishingly small.)

Watts-Strogatz Model: Degree distribution

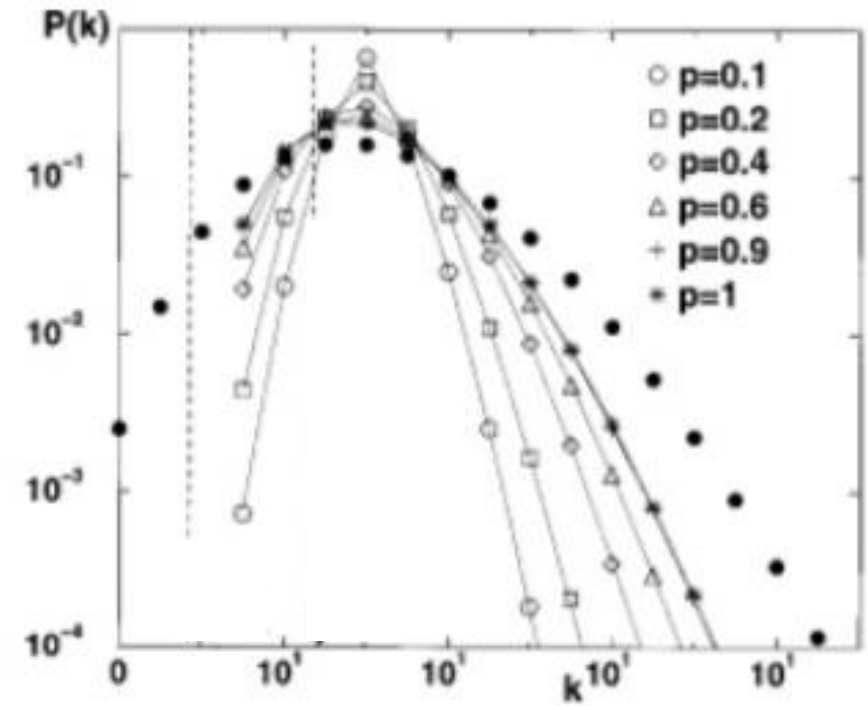

Figure 8: Degree distribution of the WS model with $N = 1000$, and various p. The filled symbols are from an equivalent random graph. Note how the WS degree dist becomes more spread out for larger p.

Watts-Strogatz model: Degree distribution

In the WS model for $p = 0$, each node has the same degree K. This appears as a delta in the degree distribution function at $k = K$.

A non-zero p introduces some disorder and broadens the degree distribution $P(k)$ away from K.

When rewiring, we rewire only one end of an edge (a total of $pNK/2$ rewires) leaving each node with at least $K/2$ edges after the rewiring process. Therefore, for $K > 2$, there are no isolated nodes and the network is connected.

- J. Travers and S. Milgram, “An experimental study of the small world problem,” *Sociometry*, vol. 32, pp. 425–443, 1969.
- D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, pp. 440–442, June 1998.
- A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

Ackowledgement

- Dr Neal McBride, Arista Networks